



SUBJECT AREAS:

EVOLUTION

ENVIRONMENTAL SCIENCES

RISK FACTORS

CHEMISTRY

# Water pollutant fingerprinting tracks recent industrial transfer from coastal to inland China: A case study

Weiwei Zheng<sup>1\*</sup>, Xia Wang<sup>1\*</sup>, Dajun Tian<sup>1\*</sup>, Songhui Jiang<sup>1</sup>, Melvin E. Andersen<sup>2</sup>, Gengsheng He<sup>1</sup>, M. James C. Crabbe<sup>3</sup>, Yuxin Zheng<sup>4</sup>, Yang Zhong<sup>5,6</sup> & Weidong Qu<sup>1,7</sup>

<sup>1</sup>Key Laboratory of Public Health Safety, Ministry of Education, Department of Environmental Health, School of Public Health, Fudan University, Shanghai 200032, China, <sup>2</sup>Institutes for Chemical Safety Sciences, The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina 27709, USA, <sup>3</sup>Institute of Biomedical and Environmental Science & Technology, Faculty of Creative Arts, Technologies and Science, University of Bedfordshire, Luton LU1 3JU, United Kingdom, <sup>4</sup>National Institute of Occupational Health and Poison Control, Chinese Center for Disease Control & Prevention, Beijing 100050, China, <sup>5</sup>Institute of Biodiversity Science and Institute of High Altitude Medicine, Tibet University, Lhasa 850000, China, <sup>6</sup>School of Life Sciences, Fudan University, Shanghai 200433, China, <sup>7</sup>Center of Global Health, School of Public Health, Fudan University, Shanghai 200032, China.

Received  
21 August 2012

Accepted  
14 December 2012

Published  
7 January 2013

Correspondence and requests for materials should be addressed to Y.Z. (yxzheng@163bj.com) or Y.Z. (yangzhong@fudan.edu.cn) or W.Q. (wdqu@fudan.edu.cn)

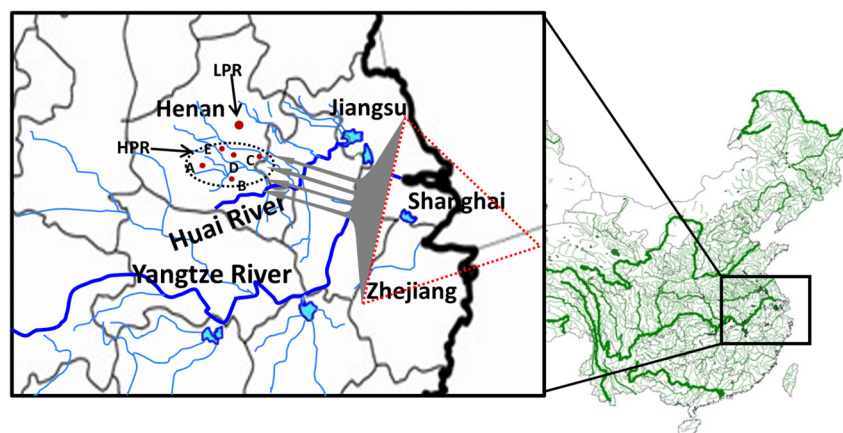
\* These authors contributed equally to this work.

In recent years, China's developed regions have transferred industries to undeveloped regions. Large numbers of unlicensed or unregistered enterprises are widespread in these undeveloped regions and they are subject to minimal regulation. Current methods for tracing industrial transfers in these areas, based on enterprise registration information or economic surveys, do not work. We have developed an analytical framework combining water fingerprinting and evolutionary analysis to trace the pollution transfer features between water sources. We collected samples in Eastern China (industrial export) and Central China (industrial acceptance) separately from two water systems. Based on the water pollutant fingerprints and evolutionary trees, we traced the pollution transfer associated with industrial transfer between the two areas. The results are consistent with four episodes of industrial transfers over the past decade. Our results also show likely types of the transferred industries - electronics, plastics, and biomedicines - that contribute to the water pollution transfer.

The 'Rise of Central China' Plan has become the cornerstone of Chinese economic development strategy<sup>1,2</sup>, following the previous development of coastal regions<sup>3</sup>. Central areas of China, with their high population densities and developmental potential, are also very attractive sites for industrial development<sup>1,4</sup>. Today, this transfer of industrial activity is already in progress between developing and developed regions. The scales, temporal-spatial pattern, transfer process and intensity of industrial transfer vary significantly from area to area<sup>5,6</sup>. Water pollution inevitably accompanies the transfer of manufacturing enterprises between regions<sup>7</sup>. Due to limited regulation and registration in the undeveloped areas, there are a large number of unlicensed or unregistered enterprises in these undeveloped regions<sup>8,9</sup>. Therefore, economic surveys or use of economic geography, the usual method for tracing industrial transfer, are not useful<sup>10,11</sup>.

We tackled the problem of accounting for pollution transfer by identifying the pollutant composition (fingerprint) in water bodies from various regions. The pollutant profiles, with appropriate analysis, reflect current pollution status of particular water sources and infer the transfer process. To assure that water fingerprint samples were representative of transfer activities, we selected two areas to serve as prototypical export and acceptance regions (Fig. 1). The export region located in the Yangtze Delta includes the developed coastal area in Eastern China<sup>1</sup>. This region has transferred various enterprises in recent years to central areas, such as Henan province<sup>4</sup>. With the soaring economic development in the past two decades, water bodies have received significant increases in organic pollutants<sup>12,13</sup>. In particular, parts of the Huai River basin located in Henan segment experienced increased organic pollution discharge<sup>12</sup>. These two areas, separately affiliated to the two main, relatively isolated river basins in China, can sufficiently visualize the image of industrial (pollution) transfer process based on their geographic, climatic, economic, developmental and urban-rural differences.

Water fingerprints evolve by a process of descent with changes<sup>14</sup>. Pollutant components of water bodies in different regions and periods originate from the same or different pollution sources; similarities among water fingerprints reflect water-body relationships (Supplementary Information 1)<sup>15</sup>. The evolutionary process gives



**Figure 1 | Industrial and pollution transfers from Yangtze Delta to Henan in China.** The elliptical broken line denotes the industrial acceptance regions belonging to the Huai River segment in Henan. A–E red dots represent sampling sites in HPR (highly polluted region); which correspond to HPRA-HPRE in Fig. 2. The other red dot indicates the sampling site of LPR (low polluted region) for the control. The triangular broken line shows a region of the Yangtze Delta, the typical industrial export region. The four grey arrows represent four times of industrial transfer from the Yangtze Delta to Henan (see Fig. 2).

rise to diversities of fingerprints among various water bodies<sup>15</sup>. Similarly, in species evolving from ancestors, random mutations and variation caused by environmental pressure (insertion, substitution and deletion) of biological sequences produce the diversities of species (Supplementary Information 1)<sup>16,17</sup>. Reconstructing ancestral origin discloses complex relationships among species based on comparison of similarity and variation between existent biological sequences<sup>16,17</sup>. Tree topology, which powerfully reflects natural history and biological phylogeny<sup>18–22</sup>, effectively identifies the relationships of the pollution fingerprints in various water bodies<sup>23</sup>.

In this study, we developed an analytical framework combining water fingerprint and evolutionary analysis, examined the diversities of water fingerprints and temporal-spatial relationships among the water sources, and explicitly traced industrial transfer processes and common features between different regions. In addition, we analyzed target compounds in the water sources for identifying possible types of industrial transfer.

## Results

Fig. 2 demonstrates the tree constructed for the water fingerprints in various regions. The red (highly polluted regions in Henan), purple (Shanghai), green (Zhejiang) and orange (Jiangsu) branches cluster together in main branch 2 and 3. The blue branches (low polluted regions in Henan) are in a separate main branch 1. Thus, fingerprints of water samples in highly polluted regions of Henan were more similar to those of Eastern regions than were those of low polluted regions (control regions) in Henan. This observation confirms the pollution transfer between the industrial export and acceptance regions. Thus, the water bodies in industrial export regions serve as a ‘source’ and those in the industrial acceptance regions appear as a ‘sink’.

In two main branches of the tree (branch 2 and branch 3), the red branches with highly polluted region A, B, C, D, and E in Henan actually clustered with the water sources from Shanghai, Zhejiang and Jiangsu in four clades - T(a), T(b), T(c), and T(d). Though A, B, C, D, and E localities had relative proximity (Fig. 1), the water fingerprint samples from these five localities were more closely related to those from water sources in Eastern regions than to each other. The local water fingerprints in A–E localities belonged to four distinct lineages. Since the reconstructed relationships of water pollution fingerprints represented the origin of pollution relationships between different regions (Supplementary Information 1), these pollution lineages are suggestive of four pollution origin episodes. Additionally, each lineage showed the same or similar pollution

origin transferred from industrial export regions to acceptance regions. Therefore, these four episodes are likely to represent independent events. T(a), T(b), T(c), and T(d) then are likely to represent four episodes of pollution (industrial) transfer from Eastern to Central China (Fig. 2).

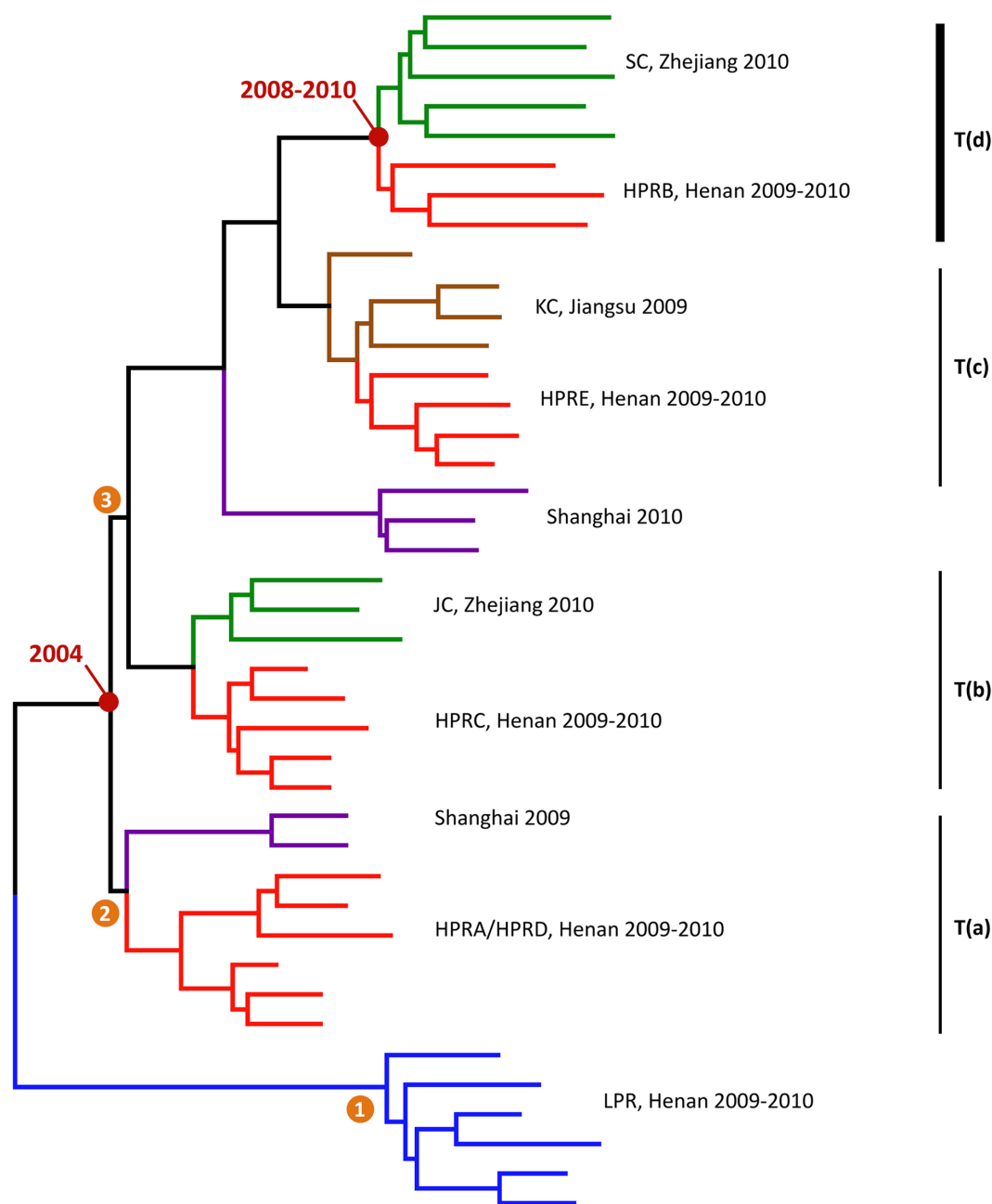
The transfer rate heterogeneity among these four pollution transfers was analyzed by the relative rate test between the lineages in the tree. T(d) had relatively faster rates than other three transfers; however, no significant rate differences were observed between every two transfers of T(a), T(b) and T(c) (Table 1 and Fig. 2). Therefore, the grouping [T(a), T(b) and T(c)] had different transfer features compared to T(d), indicating two different stages of industrial transfer from east to west.

Our analysis of the target compounds in the water samples revealed that the water fingerprints from high polluted regions in Henan and from the Yangtze Delta identified relatively high concentrations of plasticizers, polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs) and sulfonamide or steroid compounds; these pollutants were not found in the low pollution area of Henan (Table 2).

## Discussion

Based on the specific outcomes from the reconstructed ‘pollution cladogram’ and rate heterogeneity test for transfer, we now can focus on the process of these four episodes of industrial transfer. In China, the industrial transfer began in 2004<sup>24</sup>. Local economic statistics also indicated that the highly polluted regions in Henan accepted industries on a relative large scale for the first time during 2004–2005 (Supplementary Information 2). Owing to the global financial crisis after 2008<sup>4</sup>, China experienced significant adjustments and changes in industrial policies and economic strategies. During this period, the developed regions accelerated the pace of industrial transfer. Thus, we hypothesized that T(d), with its faster rate, happened in this more recent economic cycle and was the fourth transfer. The previous three industrial transfers, T(a) through T(c), occurred during 2004–2008, with a constant rate of one episode every sixteen months (Table 1).

The occurrence or evolution of pollution is now extremely rapid in this era of intensive industrialization. Our results also show that pollution transfers vary on a year-to-year scale in emerging economies. The ability to estimate timing of events in industrial transfer processes requires much smaller errors than in biological or linguistic evolution processes. Based on the inferred errors in the rate heterogeneity test (Table 1), each constant rate has an associated



**Figure 2 | Industrial transfer process reconstructed based on water pollution fingerprints.** Branch 1, 2 and 3 (the numbers labels are next to the relevant branches), colored by locality, denote the three major branches in the tree. Locations in LPR (low polluted region) in Henan are in blue; locations (HPRA-HPRE) in HRP (high polluted region) in Henan are in red. Water samples from Shanghai, Zhejiang and Jiangsu are respectively in purple, green and orange. T(a)–T(d) indicate the four estimated industrial transfers. The thicker line of T(d) represents the faster transfer rate than other three transfers. The red dots and values represent the inferred time points in the industrial transfer process.

uncertainty of  $\pm 0.2$  year (about two and a half months). This correction can provide a more precise estimation for transfer rates.

The analysis on the industrial transfer process illustrates the need to confirm the conclusions of the ‘pollution transfer tree’ by considerations of economic data and industrial cycle information. Although these demographic data are limited, they could refine ideas about the main periods of transfer and give valuable clues in tracing the industrial transfer process ongoing in Eastern China. Unlike the conclusions from biological evolutionary trees of a single origin species and subsequent descent, we did not find any single origin of water fingerprints from these various sources. Nonetheless, because local pollutant releases are relatively steady within the time-periods

of our observations, the water pollutant compositions and the whole fingerprints provide an overview of pollution features.

Specific pollutant compounds can serve as index compounds or industrial tracers for inferring the nature of industries involved in the pollution. Recycling procedures for electronic wastes (E-waste) including computers, cellular phones, and televisions<sup>25,26</sup> release PCBs and PBDEs from capacitors or transformers. With the export of computers and electronic industry wastes from developed countries, Eastern China became one of the important E-waste recycling regions with widespread distribution of these compounds in various environmental media and into the general population<sup>27</sup>. Previously, the electronic industry in Henan was undeveloped and index

**Table 1 | Results of relative rate tests for four industrial transfers**

	T(a)	T(b)	T(c)	T(d)	Estimated transfer rate
T(a)	—	0.49	1.07	4.23**	once every sixteen months
T(b)		—	−1.84	3.74**	once every sixteen months
T(c)			—	2.09**	once every sixteen months
T(d)				—	faster than once every sixteen months

T(a), T(b), T(c) and T(d) denote four times of industrial transfer estimated from the pollution origin relationship tree, see Fig. 1. The infer error of estimated rate for three times of transfer with constant speed was  $\pm 0.2$  year (about two and a half months). Each value in the table denotes  $K_{ij}/SE$ , which is given as  $[K_i - K_j]/SE$ . So that a  $K_{ij}/SE > 0$  means that the  $i$ th column is estimated to have faster transfer rate than the row.  
\*\* $P < 0.01$ .

pollutants for E-waste such as PCBs or PBDEs were uncommon in environmental media<sup>27</sup>. The transferred PCBs and PBDEs in sink type water bodies strongly suggest that the electronic industry has shifted various activities to the acceptance region.

Plasticizers appear in a wide range of products<sup>28</sup>. The plasticizers detected in the highly polluted regions of Henan indicated that the plastic industry has also transferred activities from Yangtze Delta.

Sulfonamide and steroid compounds are intermediate products or metabolites for many pharmaceuticals<sup>29</sup>. Sulfonamides are a kind of antibiotics: both animals and humans produce and excrete various steroids. In many cases, these compounds are more used as tracers for domestic or urban waste water than as industrial tracers<sup>30,31</sup>. However, in the control regions (also in Henan and population in these regions have similar lifestyle with those in highly polluted regions), the sulfonamides and steroids were not detected in water samples (Table 2). Additionally, in the upper reaches of highly polluted regions in Henan, the target sulfonamides and steroids were not detected in surface or ground waters. These observations suggested that target sulfonamides and steroids may also act as industrial tracers or index compounds in this particular case. Therefore, the detected sulfonamides and steroids in water samples implied that biomedical industries were also accepted by local areas in Henan.

These categories of industry - E-waste recycling, plastics, and biomedicines - all previously existed in Henan. However, the detection of these compounds at relatively high levels in sink type water bodies suggested that these types of industry had expanded their presence in acceptance localities in Henan.

Industrial enterprises, with their associated pollution, have transferred from the developed regions and are invading inland rural areas in China. This demographic shift is inducing serious pressures on vulnerable rural ecology, the environment and public health. Our pollution-transfer-tree unraveled the complexities of the industrial transfer process and provided results to enhance the scientific bases for pollution control and environmental protection.

Based on projections of growth and movement of industries westward, we can predict the industrial transfer trends in the future. For

instance, it was estimated that Henan will experience unprecedented opportunities for development during the 12th Five-Year Plan (2011–2015)<sup>32</sup>. Based on this observation of faster transfer after 2008, we expect more than three times the rate of industrial (pollution) transfer wave from Yangtze Delta to Henan in the coming four years. Another value of our tree analysis is the ability to trace industrial transfer processes and uncover transferred industry types. For example, two kinds of industrial transfer process can be traced separately: from developed countries to developed regions in China and from China's developed regions to undeveloped regions. Analyzing results in this manner can give valuable information for comparing differences of types, dynamics, effects and patterns between intra- and inter-country pollutant transfers.

The government could acquire sufficient pollution transfer information by monitoring of water pollution fingerprints on a regular and long-term manner covering more regions. Such a process would outline a more integrated and macro- image for industrial transfer over a longer time scale within the whole country, even inter-countries independently from the economic survey. This expanded monitoring can provide a supplement to the economic survey, and even play dominant roles if the monitoring data are more solid and convincing.

China, suffering from regional imbalance and experiencing unprecedented development in multiple geographic regions, provides a unique opportunity to evaluate pollution consequences of industrialization. Our pollution tree framework is generic. Other regions in China and developing countries facing similar challenges trace pollution transfer and monitor development through use of these tools.

## Methods

**Sampling strategy.** We selected Shanghai, Zhejiang and Jiangsu provinces as study sites in the Yangtze Delta, the most typical and important industrial export regions in China. Water samples from the most representative water plants in the developed cities in Yangtze Delta were collected, as they can show the pollution fingerprints of urban water sources and drinking water. Henan province, one of the important regions accepting industries from the Yangtze Delta and experiencing soaring economic development in recent decades, was selected. We selected five localities in highly polluted regions of Henan for water sampling (A, B, C, D, and E; Fig. 1). These localities have significantly more industrial enterprises from transfer than from local investment or foreign countries (see Supplementary Information 2). The water samples were collected from the main rivers receiving the discharge of aggregated enterprises along the river. The ground water samples were collected in the villages neighboring the polluted rivers. To examine and distinguish the pollution in water from historical pollution discharge or from recent period of industrial transfer, we also selected low pollution regions as a historical 'control regions' in Henan (Fig. 1 and Supplementary Table S1). These 'control regions' were distant from the tributary of the Huai River. Since no irrigation canals with water sourced from polluted regions flow through this area, the river basin in this area has no connection to highly polluted river basin. Additionally, in the sampling drainage area, there are less than 20 enterprises on average (Supplementary Table S1).

**Water samples.** In all the selected sites, a total of 41 water samples were collected during 2009–2010 to cover the main types of local water bodies (Supplementary Table S2). Based on drainage basin features and surrounding industrial enterprises, we respectively sampled surface and ground water in highly polluted and low polluted areas (control areas) of Henan. In Shanghai, Zhejiang and Jiangsu, we sampled water source, raw water and drinking water from the municipal water plants.

**Table 2 | Quantitative results of main target compounds in the water fingerprints ( $\mu\text{g/L}$ )**

	Yangtze Delta	Highly polluted localities in Henan	Control regions in Henan
Plasticizers	0.07–12.6	0.4–16.1	0.005–0.012
PCBs	0.01–0.07	0.02–0.08	ND <sup>a</sup>
PBDEs	ND–32.4 <sup>a,b</sup>	ND–28.7 <sup>a,b</sup>	ND <sup>a</sup>
Sulfonamides	ND–0.09 <sup>a</sup>	ND–0.03 <sup>a</sup>	ND <sup>a</sup>
Steroids	ND–5.6 <sup>a,c</sup>	ND–2.1 <sup>a,c</sup>	ND <sup>a</sup>

<sup>a</sup>None detected;  
<sup>b</sup>unit: pg/L;  
<sup>c</sup>unit: ng/L.





**Whole-pollution-fingerprints analysis.** To examine pollutant discharge, organic fingerprints of various water samples, analyzed by gas chromatography-mass spectrum (GC-MS) and liquid chromatography-tandem mass spectrometry (LC-MS/MS), traced the pollutant discharges and provided the data matrix for creating the evolutionary tree. Each water sample (40 L) was first filtered through a 0.45  $\mu\text{m}$  glass fiber membrane and extracted by passing through 10 g XAD-2 cartridge. After elution and concentration, the whole pollutant fingerprint in each extract was detected by GC-MS<sup>23,33</sup> and LC-MS/MS (detailed information in Supplementary Information 4 and Supplementary Information 5).

**Common compound matching between water fingerprints.** We developed programs and macrocommands in SAS 9.2<sup>34</sup> for automatically matching the resolved spectra between every two water pollution samples. As the retention time of the peaks could not give sufficient information for identifying compounds, the resolved mass spectra were combined to ascertain that the same compound was represented by the same variable number in all samples. We evaluated similarity between spectra depending on several guides<sup>23,35,36</sup>. Finally, we calculated the integrated areas of the remaining resolved chromatograms. In the resulting data matrix (X), each row represented one sample and each column denoted one compound, the latter identified by its mean retention time. The numbers of detected peaks in all water fingerprints from 41 water samples ranged from 240 to 392. The total number of common compounds (appeared in at least two samples) in all water pollution fingerprints was 1522 based on similarity matching among all resolved peaks using both retention time and mass spectrum information. On average, between every two pollution fingerprints of water samples, 148 common compounds (ranging from 115 to 192) were matched.

**Data matrix conversion<sup>23</sup>.** After establishing a data matrix (X), the presence or absence of each compound was coded as '1' or '0', respectively, to produce a binary matrix (Y) of all matched and included compounds in X. Then the converted data matrix (Y) can be used for character-feature-based tree-building methods.

**Distance matrix establishment<sup>23</sup>.** Using the information of overlapped compounds between every two water fingerprints, we defined and calculated a distance matrix of 'intersection and union ratio (IUR)' of all 41 water pollution samples. Detailed descriptions and formulae are in Supplementary Information 3 and Supplementary Table S3.

**Tree construction.** The neighbor-joining (NJ) method was used to construct trees of complex water fingerprints through MEGA 4.1<sup>37</sup> based on IUR distance matrix. The binary data matrix (Y) was used for constructing 'pollution trees' by maximum parsimony (MP) and maximum likelihood (ML) methods through PAUP 4.0<sup>38</sup>. Topological robustness was investigated using 10000 non-parametric bootstrap replicates.

**Relative rate test.** The relative-rate tests of between the lineages in the water fingerprint tree were performed with the proposed method<sup>39</sup> and using PAML 4.5<sup>40</sup>. The binary transferred matrix was used for relative rate tests.

**Quantitative analysis for target compounds.** The target compounds in the water fingerprints such as plastics, PCBs, PBDEs, sulfonamides and steroidal compounds (the detailed compound list see Supplementary Table S4) were quantitatively investigated. The phthalate esters (PAEs), PCBs and PBDEs were detected by solid phase extraction (SPE) with GC-MS<sup>41–44</sup>. To detect the trace-level PCBs and PBDEs in water samples, a large-volume (20 L) of water was used for extracting target compounds in each sample<sup>41,44</sup>. Sulfonamides and steroids were analyzed by SPE with liquid chromatography-tandem mass spectrometry (LC-MS/MS)<sup>41,45</sup>. Only the sulfonamides were quantified by external standard curves, all the other target compounds were quantified by internal standard calibration procedures. The correlation coefficients of the calibration curves for all target compounds were greater than 0.99 (Supplementary Information 5 and Supplementary Table S5, S7, S8 and S9). The limit of detection (LOD) was defined as a signal of 3 times the noise level. The LOD for PAEs (0.2–1.4 ng/L), and PCBs (2–8 pg/L), PBDEs (0.08–6 pg/L), sulfonamides (0.2–0.8 ng/L) and steroids (0.02–0.06 ng/L) are shown in Supplementary Information 5 and Supplementary Table S5, S7, S8 and S9. For every batch of 10 samples, a solvent blank and a procedural blank were added. The recoveries of surrogate standards, spiked blank samples and spiked matrix blank samples were analyzed to evaluate the repeatability and accuracy of analytical procedures. The sample preparation, parameters of instrumental analysis and results of quality assurance (QA) and quality control (QC) are shown in Supplementary Information (Supplementary Information 5 and Supplementary Table S5–S9).

- Zheng, Y. N. & Chen, M. J. China's regional disparity and its policy response. <http://www.nottingham.ac.uk/cpi/documents/briefings/briefing-25-china-regional-disparity.pdf> (2007) (Accessed 15th March 2011).
- Yu, Y. Y. & Gao, L. Central China rise and ideological construction. *Int. Biz. Mgmt.* **1**, 1–6 (2010).
- Ke, S. Agglomeration, productivity, and spatial spillovers across Chinese cities. *Ann. Reg. Sci.* **45**, 157–179 (2010).

- Liu, D. Factor flow and transfer of the processing trade in China: from coastline to central and west areas. <www.academicpub.org/essr/download.aspx?ID=118> (2011) (Accessed 20th January 2012).
- Liu, Y. & Huang, Y. Impact analysis of international industrial transfer on SMEs in Zhejiang. *2011 International Conference on E-business and E-government*. <http://ieeexplore.ieee.org/xpls/abs\_all.jsp?arnumber=5882151> (2011) (Accessed 20th January 2012).
- Huai, J. Industrial transfer and economic development in less developed regions. *2010 International Conference on E-business and E-government*. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=5592344> (2010) (Accessed 2nd December 2011).
- Mehra, M. K. & Das, S. P. North-South trade and pollution migration: the debate revisited. *Environ. Resources Econ.* **40**, 139–164 (2007).
- State Administration for Industry & Commerce of the People's Republic of China. *Yearbook of Industry and Commerce Administration of China 2011*. (China Industry and Commerce Associated Press, Beijing, 2011) (in Chinese).
- Kushnir, K., Mirmulstein, M. L. & Ramalho, R. Micro, small, and medium enterprises around the world: how many are there, and what affects the count? IFC and the World Bank. <http://www1.ifc.org/wps/wcm/connect/9ae1dd80495860d6a482b519583b6d16/MSME-CI-AnalysisNote.pdf?MOD=AJPERES&CACHEID=9ae1dd80495860d6a482b519583b6d16> (2010) (Accessed 2nd December 2011).
- Boschma, R. & Frenken, K. Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *J. Econ. Geogr.* **6**, 273–302 (2006).
- Frenken, K. & Boschma, R. A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. *J. Econ. Geogr.* **7**, 635–649 (2007).
- National Bureau of Statistics of China. *China Statistical Yearbook 2001–2010*. <http://www.stats.gov.cn/tjsj/ndsj/> (2001–2010) (Accessed 6th June 2011).
- Bao, L. J., Maruya, K. A., Snyder, S. A. & Zeng, E. Y. China's water pollution by persistent organic pollutants. *Environ. Pollut.* **163**, 100–108 (2012).
- Qureshi, I. H. Antarctic marine sediments as fingerprints of pollution migration. *J. Radioanal. Nucl. Ch.* **250**, 97–107 (2001).
- Pal, A., Gin, Y. H., Lin, Y. C. & Reinhard, M. Impacts of emerging organic contaminants on freshwater resources: Review of recent occurrences, sources, fate and effects. *Sci. Total Environ.* **408**, 6062–6069 (2010).
- Nei, M. & Kumar, S. *Molecular evolution and phylogenetics*. (Oxford Press, New York, 2000).
- Yang, Z. *Computational molecular evolution*. (Oxford Univ. Press, 2006).
- Barbrook, A. C., Howe, C. J., Blake, N. & Robinson, P. The phylogeny of The Canterbury Tales. *Nature* **394**, 839 (1998).
- Gray, R. D. & Atkinson, Q. D. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* **426**, 435–439 (2003).
- Rexová, K., Frynta, D. & Zrzavý, J. Cladistic analysis of languages: Indo-European classification based on lexicostatistical data. *Cladistics* **19**, 120–127 (2003).
- Pagel, M., Atkinson, Q. D. & Meade, A. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* **449**, 717–720 (2007).
- Dunn, M., Greenhill, S. J., Levinson, S. C. & Gray, R. D. Evolved structure of language shows lineage-specific trends in word-order universal. *Nature* **473**, 79–82 (2011).
- Zheng, W. *et al.* Pollution trees: Identifying similarities among complex pollutant mixtures in water and correlating them to mutagenicity. *Environ. Sci. Technol.* **46**, 7274–7282 (2012).
- People's Daily Online. Industrial transfer boosts central China. <http://english.peopledaily.com.cn/200604/19/eng20060419\_259440.html> (2006) (Accessed 25th October 2011).
- Ni, H. G., Zeng, H., Tao, S. & Zeng, E. Y. Environmental and human exposure to persistent halogenated compounds derived from e-waste in China. *Environ. Toxicol. Chem.* **29**, 1237–47 (2010).
- Robinson, B. H. E-waste: an assessment of global production and environmental impacts. *Sci. Total Environ.* **408**, 183–91 (2009).
- Breivik, K., Gioia, R., Chakraborty, P., Zhang, G. & Jones, K. C. Are reductions in industrial organic contaminants emissions in rich countries achieved partly by export of toxic wastes? *Environ. Sci. Technol.* **45**, 9154–9160 (2011).
- Zhou, Y., Wang, H., Chen, Y. & Jiang, Q. Environmental and food contamination with plasticisers in China. *Lancet* **378**, e4 (2011).
- Peng, X., Tan, J., Tang, C., Yu, Y. & Wang, Z. Multiresidue determination of fluoroquinolone, sulfonamide, trimethoprim, and chloramphenicol antibiotics in urban waters in China. *Environ. Toxicol. Chem.* **27**, 73–79 (2008).
- Li, W., Shi, Y., Gao, Y., Liu, J. & Cai, Y. Occurrence of antibiotics in water, sediments, aquatic plants, and animals from Baiyangdian Lake in North China. *Chemosphere* **89**, 1307–1315 (2012).
- Tölgyesi, A., Verebey, Z., Sharma, V. K., Kovacsics, L. & Fekete, J. Simultaneous determination of corticosteroids, androgens, and progesterone in river water by liquid chromatography-tandem mass spectrometry. *Chemosphere* **78**, 972–979 (2010).
- People's Daily Online. Henan's advantages go hand in hand with development. <http://www.chinadaily.com.cn/kindle/2011-09/08/content\_13647469.htm> (2011) (Accessed 18th February 2012).



33. Chen, L. *et al.* Status of trace organic pollution in the network water came from Huangpu River. *J. Hygiene. Res.* **3**, 137–143 (in Chinese) (2008).
34. SAS Institute, Cary, North Carolina (2008).
35. Eide, I. *et al.* Resolution of GC-MS data of complex PAC mixtures and regression modeling of mutagenicity by PLS. *Environ. Sci. Technol.* **35**, 2314–2318 (2001).
36. Eide, I., Neverdal, G., Thorvaldsen, B., Grung, B. & Kvalheim, O. M. Toxicological evaluation of complex mixtures by pattern recognition: correlating chemical fingerprints to mutagenicity. *Environ. Health. Perspect.* **110**(suppl 6), 985–988 (2002).
37. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA, version 4.1. <<http://www.megasoftware.net/mega4/index.html>> (2008) (Accessed 12th April 2010).
38. Swofford, D. L. PAUP\*. Phylogenetic Analysis Using Parsimony (\* and Other Methods). Version 4. (Sinauer Associates: Sunderland, Massachusetts, 2002).
39. Li, W. H. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* **36**, 96–99 (1993).
40. Yang, Z. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol. Bio. Evol.* **24**, 1586–1591 (2007).
41. Dachs, J. & Bayona, J. M. Large volume preconcentration of dissolved hydrocarbons and polychlorinated biphenyls from sea water: Intercomparison between C<sub>18</sub> disks and XAD-2 column. *Chemosphere* **35**, 1669–1679 (1997).
42. Berrojalbiz, N. *et al.* Persistent organic pollutants in Mediterranean seawater and processes affecting their accumulation in plankton. *Environ. Sci. Technol.* **45**, 4315–4322 (2011).
43. Mai, B. *et al.* Distribution of polybrominated diphenyl ethers in sediments of the Pearl River Delta and adjacent South China Sea. *Environ. Sci. Technol.* **39**, 3521–3527 (2005).
44. Peng, X. *et al.* Concentrations, transport, fate, and releases of polybrominated diphenyl ethers in sewage treatment plants in the Pearl River Delta, South China. *Environ. Int.* **35**, 303–309 (2009).
45. Zhang, D., Lin, L., Luo, Z., Yan, C. & Zhang, X. Occurrence of selected antibiotics in Jiulongjiang River in various seasons, South China. *J. Environ. Monit.* **13**, 1953–1960 (2011).

## Acknowledgements

We would like to thank Guozhou Zhang, Lihong Yin, Wei Xie, Zhihong Zhang, Li Chen, Xiao Wei, Ying Zhou, Hao Zhang, Li Liu, and Xin Chen for the help in sample collection; Weidong Tian (Fudan University) for the advice and comments. This project was supported by National Key Technology R&D Program in the 11th Five Year Plan (No. 2006BAI19B02), Key Project of National High-tech R&D Program of China (863 Program) (2008AA062501 & 2013AA065204), National Natural Science Foundation of China (30972438, 81202165 & 81273035), Shanghai Municipal Health Bureau Leading Academic Discipline Project (No. 08GWD14), 'Dawn' Program of Shanghai Education Commission (No. 07SG01).

## Author Contributions

All authors contributed to project led by W. Qu and Y. Zhong. W. Qu, W. Zheng, Y. Zhong and Y. Zhong designed this research and analytical approach and wrote the manuscript. W. Zheng, X. Wang, D. Tian, S. Jiang and G. He collected all the water samples, analyzed water fingerprints, and established database. W. Zheng, Y. Zhong, M. E. Andersen, M. J. C. Crabbe, Y. Zhong and W. Qu conducted data analysis and wrote the manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

**How to cite this article:** Zheng, W. *et al.* Water pollutant fingerprinting tracks recent industrial transfer from coastal to inland China: A case study. *Sci. Rep.* **3**, 1031; DOI:10.1038/srep01031 (2013).